Credit Risk Analysis – HW 20 Write Up

Overview

The purpose of this model take our lending data and to preform a predictive analysis on it to determine if we can establish a model for lending to potential borrowers. A few factors takes into consideration factors such as loan size, interest rate, debt-to-income, and the total debt for borrowers. The business use of this would be to reduce the potential workload of the credit risk/underwriting team that is approving loans and give them a baseline model for a borrower that would be easier to approve/deny based on the predetermined factors of this model.

- **Accuracy**: This model is accurate with a weighted accuracy of .99 or 99%, producing an effective an actionable result based off the previously mentioned factors.
- **Precision**: This is once section that the model could use room for improvement with a precision score of .87 or 87% for the positive and a perfect 1.0 or 100% for the negative, this could suggest overfitting for the model.
- **Recall**: The recall score shows a 1.0 or 100% for the negative and a .95 or 95% for the positive meaning that it does have a high rate of determining positive instance in the data set but is not perfect with the positive score being only .95.

Summary

This machine learning model presents a specific use case for the credit risk teams ability to streamline their approval and denial of loans. The accuracy, precision and recall scores suggest that while the model is effective at determining a accurate positive or negative result, there is overfitting within the model which would weaken the recommendation for use. Additionally, the model is not scaled, we could potentially achieve a higher rate of precision and higher recall score if additional models are explored for this use case. As the model stands we can recommend it for use in production with a noted consideration. The credit analysis team must be vigilant in use with this model and apply a level of credit analysis reasoning on the results of the model and understand the input of the model prior to relying on its sole decision. While accurate, false positives and negatives exist within the dataset and may lead to consumers being denied or approved for loans they would otherwise not qualify for.